Large Deviations for Stochastic Evolution Equations with Small Multiplicative Noise *

Wei Liu †

Fakultät Für Mathematik, Universität Bielefeld, D-33501 Bielefeld, Germany

School of Mathematical Sciences, Beijing Normal University, Beijing 100875, China

Abstract

The Freidlin-Wentzell large deviation principle is established for the distributions of stochastic evolution equations with general monotone drift and small multiplicative noise. As examples, the main results are applied to derive the large deviation principle for different types of SPDE such as stochastic reaction-diffusion equations, stochastic porous media equations and fast diffusion equations, and the stochastic p-Laplace equation in Hilbert space. The weak convergence approach is employed in the proof to establish the Laplace principle, which is equivalent to the large deviation principle in our framework.

AMS subject Classification: 60F10, 60H15.
Keywords: Stochastic evolution equation, large deviation principle, Laplace principle, variational approach, weak convergence approach, reaction-diffusion equations, porous media equations, fast diffusion equations, p-Laplace equation.

1 Introduction

There mainly exist three different approaches to analyze stochastic partial differential equations (SPDE) in the literature. The "martingale measure approach" was initiated by J. Walsh in [38]. The "variational approach" was first used by Bensoussan and Temam in [3, 4] to study SPDE with additive noise, later this approach was further developed

*Supported in part by the DFG through the Internationales Graduiertenkolleg "Stochastics and Real World Models" and NNSFC(10721091).

†E-mail: weiliu0402@yahoo.com.cn Tel:+49-(0)521-1062971, Fax: +49-(0)521-1066455

in the works of Pardoux [24], Krylov and Rozovoskii [21] for more general case. For the "semigroup (or mild solution) approach" we refer to the classical monograph [12] by Da Prato and Zabcyzk. In this paper we use the variational approach to treat a large class of nonlinear SPDE of evolutionary type, which can model all kinds of dynamics with stochastic influence in nature or man-made complex systems. Stochastic evolution equations have been studied intensively in recent years and we refer to [10, 11, 20, 22, 23, 27, 30, 39, 42] for various generalizations and applications.

Concerning the large deviation principle (LDP), there also exist fruitful results within different frameworks of SPDE. The general large deviation principle was first formulated by Varadhan [35] in 1966. For its validity to stochastic differential equations in finite dimensional case we mainly refer to the well known Freidlin-Wentzell LDP ([19]). The same problem was also treated by Varadhan in [37] and Stroock in [34] by a different approach, which followed the large deviation theory developed by Azencott [2], Donsker-Varadhan [14] and Varadhan [35]. In the classical paper [18] Freidlin studied the large deviations for the small noise limit of stochastic reaction-diffusion equations. Subsequently, many authors have endeavored to derive the large deviations results under less and less restrictive conditions. We refer the reader to Da Prato and Zabczyk [12] and Peszat [25] (also the references therein) for the extensions to infinite dimensional diffusions or stochastic PDE under global Lipschitz condition on the nonlinear term. For the case of local Lipschitz conditions we refer to the work of Cerrai and Röckner [8] where the case of multiplicative and degenerate noise was also investigated. The LDP for semilinear parabolic equations on a Gelfand triple was studied by Chow in [9]. Recently, Röckner *et al* established the LDP in [32] for the distributions of the solution to stochastic porous media equations within the variational framework. All these papers mainly used the classical ideas of discretization approximations and the contraction principle, which was first developed by Freidlin and Wentzell. But the situation became much involved and complicated in infinite dimensional case since each type of nonlinear SPDE needs different specific techniques and estimates.

An alternative approach for LDP has been developed by Feng and Krutz in [17], which mainly used nonlinear semigroup theory and infinite dimensional Hamilton-Jacobi equation. The techniques rely on the uniqueness theory for the infinite dimensional Hamilton-Jacobi equation and some exponential tightness estimates.

In this paper we will study the large deviation principle for stochastic evolution equations with general monotone drift and multiplicative noise, which are more general than the semilinear case studied in [9] and the additive noise case in [32]. This framework covers all types of SPDE in [30, 21] such as stochastic reaction-diffusion equations, stochastic p-Laplace equation, stochastic porous media equations and fast diffusion equations. It is quite difficult to follow the classical discretization approach in the present case. The reason is many technical difficulties appear since the coefficients of SPDE in our framework live on a Gelfand triple. For example, it is very difficult to obtain some regularity (Hölder) estimate of the solution w.r.t. the time variable, which is essentially required in

the classical proof of LDP by discretization approach.

Hence we would use the stochastic control and weak convergence approach in this paper. This approach is mainly based on a variational representation formula for certain functionals of infinite dimensional Brownian Motion, which was established by Budhiraja and Dupuis in [5]. The main advantage of the weak convergence approach is that one can avoid some exponential probability estimates, which might be very difficult to derive for many infinite dimensional models. However, in the implement of weak convergence approach, there are still some technical difficulties appearing in the variational framework. The reason is the coefficients of SEE are nonlinear operators which are only well-defined via a Gelfand triple (so three spaces are involved). Hence we have to properly handle many estimates involving different spaces instead of just one single space. Some approximation techniques are also used in the proof.

The weak convergence approach has been used to study the large deviations for homeomorphism flows of non-Lipschitz SDEs by Ren and Zhang in [28], for two-dimensional stochastic Navier-Stokes equations by Sritharan and Sundar in [33] and reaction-diffusion type SPDEs by Budhiraja *et al* in [6]. For more references on this approach we may refer to [16, 29, 15].

Let us first recall some standard definitions and results from the large deviation theory. Let $\{X^\varepsilon\}$ be a family of random variables defined on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ and taking values in some Polish space E. Roughly speaking, the large deviation theory concerns itself with the exponential decay of the probability measures of certain kinds of extreme or tail events. The rate of such exponential decay is expressed by the "rate function".

Definition 1.1. (Rate function) A function $I : E \to [0, +\infty]$ is called a rate function if I is lower semicontinuous. A rate function I is called a good rate function if the level set $\{x \in E : I(x) \le K\}$ is compact for each $K < \infty$.

Definition 1.2. (Large deviation principle) The sequence $\{X^\varepsilon\}$ is said to satisfy the *large deviation principle with rate function* I if for each Borel subset A of E

$$-\inf_{x\in A^o} I(x) \le \liminf_{\varepsilon\to 0} \varepsilon^2 \log \mathbf{P}(X^\varepsilon \in A) \le \limsup_{\varepsilon\to 0} \varepsilon^2 \log \mathbf{P}(X^\varepsilon \in A) \le -\inf_{x\in \bar{A}} I(x),$$

where A^o and $\bar{A}$ are respectively the interior and the closure of A in E.

If one is interested in obtaining the exponential estimates on general functions instead of the indicator functions of Borel sets in E, then one can study the following Laplace principle (LP).

Definition 1.3. (Laplace principle) The sequence $\{X^\varepsilon\}$ is said to satisfy the *Laplace principle with rate function* I if for each bounded continuous real-valued function h defined on E

$$\lim_{\varepsilon\to 0} \varepsilon^2 \log \mathbf{E} \left\{ \exp\left[-\frac{1}{\varepsilon^2} h(X^\varepsilon) \right] \right\} = -\inf_{x\in E} \{h(x) + I(x)\}.$$

The starting point for the weak convergence approach is the equivalence between LDP and LP if E is a Polish space and the rate function is good. This result was first formulated in [26] and it is essentially a consequence of Varadhan's lemma [35] and Bryc's converse theorem [7]. We refer to [16, 13] for an elementary proof of it.

Let $\{W_t\}_{t\geq 0}$ be a cylindrical Wiener process on a separable Hilbert space U w.r.t a complete filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$ (i.e. the path of W take values in $C([0,T];U_1)$, where U_1 is another Hilbert space such that the embedding $U \subset U_1$ is Hilbert-Schmidt). Suppose $g^\varepsilon : C([0,T];U_1) \to E$ is a measurable map and $X^\varepsilon = g^\varepsilon(W_\cdot)$. Let

$$\mathcal{A} = \left\{ v : v \text{ is } U\text{-valued } \mathcal{F}_t\text{-predictable process and } \int_0^T \|v_s(\omega)\|_U^2 \mathrm{d}s < \infty \ a.s. \right\},$$

$$S_N = \left\{ \phi \in L^2([0,T],U) : \int_0^T \|\phi_s\|_U^2 \mathrm{d}s \leq N \right\}.$$

The set S_N endowed with the weak topology is a Polish space (we will always refer to the weak topology on S_N in this paper if we don't state it explicitly). Define

$$\mathcal{A}_N = \{ v \in \mathcal{A} : v(\omega) \in S_N \ \mathbf{P} - a.s. \}.$$

Now we formulate the following sufficient condition for the Laplace principle (equivalently, large deviation principle) of X^ε as $\varepsilon \to 0$.

(A) There exists a measurable map $g^0 : C([0,T];U_1) \to E$ such that the following two conditions hold:

(i) Let $\{v^\varepsilon : \varepsilon > 0\} \subset \mathcal{A}_N$ for some $N < \infty$. If v^ε converge to v in distribution as S_N-valued random elements, then

$$g^\varepsilon \left(W_\cdot + \frac{1}{\varepsilon} \int_0^\cdot v_s^\varepsilon \mathrm{d}s \right) \to g^0 \left(\int_0^\cdot v_s \mathrm{d}s \right)$$

in distribution as $\varepsilon \to 0$.

(ii) For each $N < \infty$, the set

$$K_N = \left\{ g^0 \left(\int_0^\cdot \phi_s \mathrm{d}s \right) : \phi \in S_N \right\}$$

is a compact subset of E.

Lemma 1.1. *[5, Theorem 4.4] If $X^\varepsilon = g^\varepsilon(W)$ and the assumption* (**A**) *holds, then the family $\{X^\varepsilon\}$ satisfies the Laplace principle (hence large deviation principle) on E with the good rate function I given by*

$$I(f) = \inf_{\left\{\phi \in L^2([0,T];U):\ f = g^0(\int_0^\cdot \phi_s \mathrm{d}s)\right\}} \left\{ \frac{1}{2} \int_0^T \|\phi(s)\|_U^2 \mathrm{d}s \right\}. \tag{1.1}$$

We will verify the sufficient condition **(A)** for general SPDE within the variational framework. Besides the classical monotone conditions assumed for the well-posedness of SPDE, we need to require one additional assumption (see $(A4)$ below) on the noise coefficient for the LDP. In fact, the weak convergence approach are used here to avoid the time discretization for SPDE (the most technical and difficult step in the classical proof of LDP) since the regularity estimate of the solution w.r.t. the time variable is unavailable in the variational framework. But unlike the semilinear case (e.g.[6]), we have to use Itô's formula for the square norm of the solution in the estimate. Then the weak convergence of control v^ε to v (see (i) of **(A)**) cause some technical difficulty in the proof of convergence of corresponding solutions under the variational framework. Hence we need to have some restriction on the noise (see $(A5)$) such that the weak convergence procedure can be verified. Later some standard approximation techniques are used to relax this assumption.

2 Main framework and result

Let

$$V \subset H \equiv H^* \subset V^*$$

be a Gelfand triple, i.e. V is a reflexive and separable Banach space and V^* is its dual space, $(H, \langle \cdot, \cdot \rangle_H)$ is a separable Hilbert space and identified with its dual space by Riesz isomorphism, V is continuously and densely embedded in H. The dualization between V^* and V is denoted by ${}_{V^*}\langle \cdot, \cdot \rangle_V$ and it is obvious that

$${}_{V^*}\langle u, v \rangle_V = \langle u, v \rangle_H, \ u \in H, v \in V.$$

Let $\{W_t\}_{t \geq 0}$ be a cylindrical Wiener process on a separable Hilbert space U w.r.t a complete filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbf{P})$. $(L_2(U; H) \| \cdot \|_2)$ denote the space of all Hilbert-Schmidt operators from U to H. We use $L(X, Y)$ to denote the space of all bounded linear operators from space X to Y.

Consider the following stochastic evolution equation

$$\mathrm{d}X_t = A(t, X_t)\mathrm{d}t + B(t, X_t)\mathrm{d}W_t, \tag{2.1}$$

where $A : [0, T] \times V \to V^*$ and $B : [0, T] \times V \to L_2(U; H)$ are measurable. For the large deviation principle we need to assume the following conditions, which are slightly stronger than those assumed in [21] for the existence and uniqueness of strong solution to (2.1).

For a fixed $\alpha > 1$, there exist constants $\delta > 0$ and K such that the following conditions hold for all $v, v_1, v_2 \in V$ and $t \in [0, T]$.

$(A1)$ (Hemicontinuity) The map $s \mapsto {}_{V^*}\langle A(t, v_1 + sv_2), v \rangle_V$ is continuous on $\mathbb{R}$.

$(A2)$ (Strong monotonicity)

$$2_{V^*}\langle A(t,v_1)-A(t,v_2),v_1-v_2\rangle_V+\|B(t,v_1)-B(t,v_2)\|_2^2\le-\delta\|v_1-v_2\|_V^\alpha+K\|v_1-v_2\|_H^2.$$

$(A3)$ (Boundedness) $\sup_{t\in[0,T]}\|B(t,0)\|_2<\infty$ and

$$\|A(t,v)\|_{V^*}+\|B(t,v)\|_{L(U,V^*)}\le K(1+\|v\|_V^{\alpha-1}).$$

$(A4)$ Suppose there exist a sequence of subspaces $\{H_n\}$ such that

$$H_n\subseteq H_{n+1},\ H_n\hookrightarrow V\text{compact and }\bigcup_{n=1}^{\infty}H_n\subseteq H\text{ dense},$$

and for any $M>0$

$$\sup_{(t,v)\in[0,T]\times S_M}\|P_nB(t,v)-B(t,v)\|_2\to0\ (n\to\infty),\tag{2.2}$$

where $P_n:H\to H_n$ is the projection operator and $S_M=\{v\in V:\|v\|_H\le M\}$.

Remark 2.1. (i) By $(A2)$ and $(A3)$ we can easily obtain the coercivity and boundedness of A and B:

$$2_{V^*}\langle A(t,v),v\rangle_V+\|B(t,v)\|_2^2+\frac{\delta}{2}\|v\|_V^\alpha\le C(1+\|v\|_H^2),$$
$$\|B(t,v)\|_2^2\le C(1+\|v\|_H^2+\|v\|_V^\alpha).$$

Hence the boundedness of B in $(A3)$ automatically holds if $\alpha\ge2$. If $1<\alpha<2$, the additional assumption on B in $(A3)$ is assumed for the well-posedness of the skeleton equation (see (2.5)).

(ii) Since for all $(t,v)\in[0,T]\times V$ we have

$$\|P_nB(t,v)-B(t,v)\|_2\to0\ (n\to\infty).$$

Hence a simple sufficient condition for (2.2) holds is to assume that

$$\{B(t,v):(t,v)\in[0,T]\times S_M\}$$

is a relatively compact set in $L_2(U;H)$. For example, we can take

$$B(t,v)=\sum_{i=1}^N b_i(v)B_i(t),$$

where $b_i(\cdot):V\to\mathbb{R}$ are Lipschitz functions and $B_i(\cdot):[0,T]\to L_2(U;H)$ are continuous.

Another simple example is $B(t,v) = QB_0(t,v)$ where $Q \in L_2(H;H)$ and

$$B_0 : [0,T] \times V \to L(U;H), \quad \sup_{(t,v)\in[0,T]\times S_M} \|B_0(t,v)\|_{L(U;H)} < \infty, \ \forall M > 0.$$

(iii) If there exists a Hilbert space H_0 such that the embedding $H_0 \subseteq H$ is compact, $\{e_i\} \subseteq H_0 \cap V$ is an ONB in H_0 and also orthogonal in H. Suppose for all $M > 0$

$$\sup_{(t,v)\in[0,T]\times S_M} \|B(t,v)\|_{L_2(U;H_0)} < \infty.$$

Then (2.2) holds. Because $B(t,v) = \sum_{i,j=1}^{\infty} b_{i,j}(t,v) u_i \otimes e_j$, by assumptions we know $\|e_j\|_H^2 \to 0$ and

$$\sup_{(t,v)\in[0,T]\times S_M} \sum_{i,j=1}^{\infty} b_{i,j}^2(t,v) < \infty.$$

then

$$\|P_n B(t,v) - B(t,v)\|_2^2 = \sum_{i=1}^{\infty} \sum_{j=n+1}^{\infty} b_{i,j}^2(t,v) \|e_j\|_H^2.$$

Hence (2.2) follows from the dominated convergence theorem. □

If $(A1) - (A3)$ hold, according to [21, Theorem II2.1] for any $X_0 \in L^2(\Omega \to H; \mathcal{F}_0; \mathbf{P})$ (2.1) has an unique solution $\{X_t\}_{t\in[0,T]}$ which is an adapted continuous process on H such that $\mathbf{E}\int_0^T (\|X_t\|_V^\alpha + \|X_t\|_H^2)\,\mathrm{d}t < \infty$ and

$$\langle X_t, v\rangle_H = \langle X_0, v\rangle_H + \int_0^t {}_{V^*}\langle A(s,X_s), v\rangle_V \mathrm{d}s + \int_0^t \langle B(s,X_s)\mathrm{d}W_s, v\rangle_H, \ \mathbf{P} - a.s.$$

holds for all $v \in V$ and $t \in [0,T]$. Moreover, we have $\mathbf{E}\sup_{t\in[0,T]} \|X_t\|_H^2 < \infty$ and the crucial Itô formula

$$\|X_t\|_H^2 = \|X_0\|_H^2 + \int_0^t \left(2_{V^*}\langle A(s,X_s), X_s\rangle_V + \|B(s,X_s)\|_2^2\right) \mathrm{d}s + 2\int_0^t \langle X_s, B(s,X_s)\mathrm{d}W_s\rangle_H.$$

Let us consider the general stochastic evolution equation with small noise:

$$\mathrm{d}X_t^\varepsilon = A(t,X_t^\varepsilon)\mathrm{d}t + \varepsilon B(t,X_t^\varepsilon)\mathrm{d}W_t, \quad \varepsilon > 0, \ X_0^\varepsilon = x \in H. \tag{2.3}$$

Hence the unique strong solution $\{X^\varepsilon\}$ of (2.3) takes values in $C([0,T];H) \cap L^\alpha([0,T];V)$. It's well-known that $(C([0,T];H) \cap L^\alpha([0,T];V), \rho)$ is a Polish space with the following metric

$$\rho(f,g) := \sup_{t\in[0,T]} \|f_t - g_t\|_H + \left(\int_0^T \|f_t - g_t\|_V^\alpha \mathrm{d}t\right)^{\frac{1}{\alpha}}. \tag{2.4}$$

It follows (from infinite dimensional version of Yamada-Watanabe theorem in [31]) that there exists a Borel-measurable function

$$g^\varepsilon : C([0,T];U_1) \to C([0,T];H) \cap L^\alpha([0,T];V)$$

such that $X^\varepsilon = g^\varepsilon(W)$ *a.s.*. To state our main result, let us introduce the skeleton equation associated to (2.3):

$$\frac{\mathrm{d}z_t^\phi}{\mathrm{d}t} = A(t,z_t^\phi) + B(t,z_t^\phi)\phi_t, \quad z_0^\phi = x,\ \phi \in L^2([0,T];U). \tag{2.5}$$

An element $z^\phi \in C([0,T];H) \cap L^\alpha([0,T];V)$ is called a solution to (2.5) if for any $v \in V$

$$\langle z_t^\phi, v\rangle_H = \langle x, v\rangle_H + \int_0^t {}_{V^*}\langle A(s,z_s^\phi) + B(s,z_s^\phi)\phi_s, v\rangle_V \mathrm{d}s, \quad t \in [0,T]. \tag{2.6}$$

We will prove (see Lemma 3.1) that $(A1)-(A3)$ also imply the existence and uniqueness of the solution to (2.5) for any $\phi \in L^2([0,T];U)$.

Define $g^0 : C([0,T];U_1) \to C([0,T];H) \cap L^\alpha([0,T];V)$ by

$$g^0(h) := \begin{cases} z^\phi, & \text{if } \ h = \int_0^\cdot \phi_s \mathrm{d}s \ \text{ for some } \phi \in L^2([0,T];U); \\ 0, & \text{otherwise.} \end{cases}$$

Then it's obvious that the rate function in (1.1) can be written as

$$I(z) = \inf\left\{ \frac{1}{2}\int_0^T \|\phi_s\|_U^2 \mathrm{d}s : \ z = z^\phi,\ \phi \in L^2([0,T],U) \right\}, \tag{2.7}$$

where $z \in C([0,T];H) \cap L^\alpha([0,T];V)$.

Now we formulate the main result which is a Freidlin-Wentzell type estimate.

Theorem 2.1. *Assume* $(A1)-(A4)$ *hold. For each* $\varepsilon > 0$, *let* $X^\varepsilon = \{X_t^\varepsilon\}_{t\in[0,T]}$ *be the solution to (2.3). Then as* $\varepsilon \to 0$, $\{X^\varepsilon\}$ *satisfies the LDP on* $C([0,T];H) \cap L^\alpha([0,T];V)$ *with the good rate function* I *which is given by (2.7).*

Remark 2.2. (i) According to [6, Theorem 5], we can also prove uniform Laplace principle by using the same arguments but with more cumbersome notation.

(ii) This theorem can not be applied to stochastic fast-diffusion equations in [23, 27] since $(A2)$ fails to satisfy. However, if we replace $(A2)$ by the classical monotone and coercive conditions in [21]

$(A2')$

$$2{}_{V^*}\langle A(t,v_1) - A(t,v_2), v_1 - v_2\rangle_V + \|B(t,v_1) - B(t,v_2)\|_2^2 \le K\|v_1 - v_2\|_H^2,$$
$$2{}_{V^*}\langle A(t,v), v\rangle_V + \|B(t,v)\|_2^2 + \delta\|v\|_V^\alpha \le K(1 + \|v\|_H^2).$$

Then the LDP can be established on $C([0,T];H)$ by the similar and simpler argument.

Theorem 2.2. *Assume $(A1),(A2'),(A3)-(A4)$ hold. Then as $\varepsilon \to 0$, the solution $\{X^\varepsilon\}$ of (2.3) satisfies the LDP on $C([0,T];H)$ with the good rate function I which is given by (2.7).*

Remark 2.3. Note that $(A2)$ mainly used to prove the additional convergence in $L^\alpha([0,T];V)$. Hence, if we only concern the LDP on $C([0,T];H)$, then we can prove the Theorem 2.2 under the weaker assumptions above. Since the proof is only a small modification (only consider the convergence in $C([0,T];H)$) of the argument for Theorem 2.1, we omit the details here.

The organization of the paper is as follows. In section 3, under the additional assumption $(A5)$ on B we prove Theorem 2.1 by using the weak convergence approach. Section 4 is devoted to relax the assumption $(A5)$ by some standard approximation techniques. In section 5 we apply the main results to different class of SPDEs in Hilbert space as applications.

3 Proof of Theorem 2.1 under additional assumption

In order to verify the sufficient conditions **(A)**, we need to first consider the finite dimensional noise, i.e. we approximate the diffusion coefficient B by P_nB. But for the simplicity of the notation, we formulate the following additional assumption on B:

$(A5)$ $B:[0,T]\times V \to L(U;V_0)$ satisfies

$$\|B(t,v)\|^2_{L(U;V_0)} \le C(1+\|v\|_V^\alpha + \|v\|_H^2),$$

where $V_0 \subseteq V$ is compact embedding and $C>0$ is a constant.

For the reader's convenience, we recall two well-known inequalities which used quite often in the proof. Throughout the paper, the generic constants may be different from line to line. If it is essential, we will write the dependence of the constant on parameters explicitly.

Young's inequality: Given $p,q>1$ satisfy $\frac{1}{p}+\frac{1}{q}=1$, then for any positive number σ, a, b we have

$$ab \le \sigma\frac{a^p}{p} + \sigma^{-\frac{q}{p}}\frac{b^q}{q}.$$

Gronwall's lemma: Let $F,\Phi,\Psi:[0,T]\to\mathbb{R}^+$ be Lebesgue measurable. Suppose Ψ is locally integrable and $\int_0^T \Psi(s)F(s)\mathrm{d}s<\infty$. If

$$\begin{aligned} &F(t) \le \Phi(t) + \int_0^t \Psi(s)F(s)\mathrm{d}s,\ t\in[0,T] \ \text{ or} \\ &\frac{\mathrm{d}}{\mathrm{d}t}F(t) \le \frac{\mathrm{d}}{\mathrm{d}t}\Phi(t) + \Psi(t)F(t),\ t\in[0,T),\ F(0)\le\Phi(0). \end{aligned} \tag{3.1}$$

Then

$$F(t) \leq \Phi(t) + \int_0^t \exp\left[\int_s^t \Psi(u)\mathrm{d}u\right] \Psi(s)\Phi(s)\mathrm{d}s,\ t \in [0,T]. \tag{3.2}$$

Lemma 3.1. *Assume* $(A1)-(A3)$ *hold. Let*

$$\|z\| := \sup_{t\in[0,T]} \|z_t\|_H^2 + \delta \int_0^T \|z_t\|_V^\alpha \mathrm{d}t$$

for $z \in C([0,T];H) \cap L^\alpha([0,T];V)$. *For all* $x \in H$ *and* $\phi \in L^2([0,T];U)$ *there exists a unique solution* z^ϕ *to (2.5) and*

$$\|z^\phi - z^\psi\| \leq \exp\left\{\int_0^T \left(K + \|\phi_t\|_U^2 + \|B(t,z_t^\psi)\|_2^2\right)\mathrm{d}t\right\} \int_0^T \|\phi_t - \psi_t\|_U^2 \mathrm{d}t \tag{3.3}$$

hold for some constant K *and all* $\phi, \psi \in L^2([0,T];U)$.

Proof. To verify the existence of the solution, we make use of [21, Theorem II.2.1]. First we assume $\phi \in L^\infty([0,T];U)$ and

$$\tilde{A}(s,v) := A(s,v) + B(s,v)\phi_s.$$

Then, due to $(A1)-(A3)$, it's easy to verify that $\tilde{A}$ satisfies Assumptions $A_i)(i=1,..,5)$ on page 1252 of [21].

(i) Hemicontinuity of $\tilde{A}$ follows from $(A1)$ and $(A2)$.

(ii) Monotonicity and coercivity of $\tilde{A}$ follows from $(A2)$ and $(A3)$.

(iii) Boundedness of $\tilde{A}$ follows from $(A3)$.

Therefore, by [21](or [41, Theorem 30.A])we know (2.5) has an unique solution.

For general $\phi \in L^2([0,T];U)$, we can find a sequence of $\phi^n \in L^\infty([0,T];U)$ such that

$$\phi_n \to \phi \text{ strongly in } L^2([0,T];U).$$

Let z^n be the unique solution to (2.5) for ϕ^n, we will show $\{z^n\}$ is a Cauchy sequence in $C([0,T];H) \cap L^\alpha([0,T];V)$. By using $(A2)$ we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\|z_t^n - z_t^m\|_H^2 =& 2{}_{V^*}\langle A(t,z_t^n) - A(t,z_t^m), z_t^n - z_t^m\rangle_V \\
&+ 2\langle B(t,z_t^n)\phi_t^n - B(t,z_t^m)\phi_t^m, z_t^n - z_t^m\rangle_H \\
\leq& 2{}_{V^*}\langle A(t,z_t^n) - A(t,z_t^m), z_t^n - z_t^m\rangle_V + \|B(t,z_t^n) - B(t,z_t^m)\|_2^2 \\
&+ \|\phi_t^n\|_U^2\|z_t^n - z_t^m\|_H^2 + 2\langle z_t^n - z_t^m, B(t,z_t^m)\phi_t^n - B(t,z_t^m)\phi_t^m\rangle_H \\
\leq& -\delta\|z_t^n - z_t^m\|_V^\alpha + (K + \|\phi_t^n\|_U^2)\|z_t^n - z_t^m\|_H^2 \\
&+ 2\|B^*(t,z_t^m)\left(z_t^n - z_t^m\right)\|_U\|\phi_t^n - \phi_t^m\|_U \\
\leq& -\delta\|z_t^n - z_t^m\|_V^\alpha + \|\phi_t^n - \phi_t^m\|_U^2 \\
&+ \left(K + \|\phi_t^n\|_U^2 + \|B(t,z_t^m)\|_2^2\right)\|z_t^n - z_t^m\|_H^2.
\end{aligned} \tag{3.4}$$

where B^* denote the adjoint operator of B and we also use the fact

$$\|B^*\|_{L(H;U)} = \|B\|_{L(U;H)} \leq \|B\|_2.$$

Then by the Gronwall lemma we have

$$\|z^n - z^m\| \leq \exp\left\{\int_0^T \left(K + \|\phi_t^n\|_U^2 + \|B(t, z_t^m)\|_2^2\right) \mathrm{d}t\right\} \int_0^T \|\phi_t^n - \phi_t^m\|_U^2 \mathrm{d}t. \tag{3.5}$$

By the similar argument we have

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\|z_t^n\|_H^2 =& 2{}_{V^*}\langle A(t, z_t^n), z_t^n\rangle_V + 2\langle B(t, z_t^n)\phi_t^n, z_t^n\rangle_H \\ \leq& -\frac{\delta}{2}\|z_t^n\|_V^\alpha + C(1 + \|z_t^n\|_H^2) + \|\phi_t^n\|_U^2\|z_t^n\|_H^2. \end{aligned} \tag{3.6}$$

Then by the Gronwall lemma and boundedness of ϕ^n in $L^2([0,T];U)$

$$\|z^n\| \leq C\exp\left\{\int_0^T \left(C + \|\phi_t^n\|_U^2\right) \mathrm{d}t\right\} \left(\|x\|_H^2 + T\right) \leq \textbf{Constant} < \infty. \tag{3.7}$$

Hence we have

$$\int_0^T \|B(t, z_t^m)\|_2^2 \mathrm{d}t \leq C\int_0^T \left(1 + \|z_t^m\|_H^2 + \|z_t^m\|_V^\alpha\right) \mathrm{d}t \leq \textbf{Constant} < \infty. \tag{3.8}$$

Combining (3.5),(3.8) and $\phi^n \to \phi$, we can conclude that $\{z^n\}$ is a Cauchy sequence in $C([0,T];H) \cap L^\alpha([0,T];V)$, and we denote the limit by z^ϕ.

Then by repeating the standard monotonicity argument(e.g.[41, Theorem 30.A]) one can show that z^ϕ is the solution of (2.5) corresponding to ϕ.

And (3.3) can be derived from (3.5). Hence the proof is complete. □

The following result shows that I defined by (2.7) is a good rate function.

Lemma 3.2. *Assume $(A1) - (A3)$ hold. For every $N < \infty$, the set*

$$K_N = \left\{g^0\left(\int_0^\cdot \phi_s \mathrm{d}s\right) : \phi \in S_N\right\}$$

is a compact subset in $C([0,T];H) \cap L^\alpha([0,T];V)$.

Proof. **Step 1**: we first assume B also satisfy $(A5)$. By definition we know

$$K_N = \left\{z^\phi : \phi \in L^2([0,T];U),\ \int_0^T \|\phi_s\|_U^2 \mathrm{d}s \leq N\right\}.$$

For any sequence $\phi^n \subset S_N$, we may assume $\phi^n \to \phi$ weakly in $L^2([0,T];U)$ since S_N is weakly compact. Denote z^n and z are the solutions of (2.5) corresponding to ϕ^n and ϕ respectively. Now it's sufficient to show $z^n \to z$ strongly in $C([0,T];H) \cap L^\alpha([0,T];V)$.

From (3.4) we have

$$\begin{aligned} &\|z_t^n - z_t\|_H^2 + \delta \int_0^t \|z_s^n - z_s\|_V^\alpha \mathrm{d}s \\ &\le \int_0^t (K + \|\phi_s^n\|_U^2)\|z_s^n - z_s\|_H^2 \mathrm{d}s + 2\int_0^t \langle z_s^n - z_s, B(s,z_s)(\phi_s^n - \phi_s)\rangle_H \mathrm{d}s. \end{aligned} \tag{3.9}$$

Define

$$h_t^n = \int_0^t B(s,z_s)(\phi_s^n - \phi_s)\mathrm{d}s.$$

By $(A5)$ and (3.8) we know $h^n \in C([0,T];V_0)$ and

$$\begin{aligned} \sup_{t\in[0,T]} \|h_t^n\|_{V_0} &\le \int_0^T \|B(s,z_s)(\phi_s^n - \phi_s)\|_{V_0} \mathrm{d}s \\ &\le \left(\int_0^T \|B(s,z_s)\|_{L(U,V_0)}^2 \mathrm{d}s\right)^{1/2} \left(\int_0^T \|\phi_s^n - \phi_s\|_U^2 \mathrm{d}s\right)^{1/2} \\ &\le \textbf{Constant} < \infty. \end{aligned} \tag{3.10}$$

Since the embedding $V_0 \subseteq V$ is compact and $\phi^n \to \phi$ weakly in $L^2([0,T];U)$, it's easy to show that $h^n \to 0$ in $C([0,T];V)$ by using the Arzèla-Ascoli theorem(also see e.g.[5, Lemma 3.2]) (more precisely, this convergence may only hold for a subsequence, but it's enough for our purpose since we may denote the convergent subsequence still by h^n). In particular, $h^n \to 0$ in $C([0,T];H) \cap L^\alpha([0,T];V)$.

Moreover the derivative (w.r.t. time variable) is given by

$$(h_s^n)' = B(s,z_s)(\phi_s^n - \phi_s).$$

As in the Lemma 3.1, we may assume $\phi^n, \phi \in L^\infty([0,T];U)$ first. Then by $(A3)$

$$\begin{aligned} \int_0^T \|(h_s^n)'\|_{V^*}^{\frac{\alpha}{\alpha-1}} \mathrm{d}s &\le \int_0^T \|B(s,z_s)(\phi_s^n - \phi_s)\|_{V^*}^{\frac{\alpha}{\alpha-1}} \mathrm{d}s \\ &\le C\int_0^T (1 + \|z_s\|_V^\alpha)\,\mathrm{d}s \\ &\le \textbf{Constant} < \infty. \end{aligned} \tag{3.11}$$

Hence $(h_\cdot^n)'$ is an element in $L^{\frac{\alpha}{\alpha-1}}([0,T];V^*)$.

By [41, Proposition 23.23] we have the following integration by parts formula

$$\langle z_t^n - z_t, h_t^n\rangle_H = \int_0^t {}_{V^*}\langle (z_s^n - z_s)', h_s^n\rangle_V \mathrm{d}s + \int_0^t {}_{V^*}\langle (h_s^n)', z_s^n - z_s\rangle_V \mathrm{d}s.$$

Hence one has

$$
\begin{aligned}
&\int_0^t \langle z_s^n - z_s, B(s,z_s)(\phi_s^n - \phi_s)\rangle_H \mathrm{d}s \\
=&\langle z_t^n - z_t, h_t^n\rangle_H - \int_0^t {}_{V^*}\langle (z_s^n - z_s)', h_s^n\rangle_V \mathrm{d}s \\
=&\langle z_t^n - z_t, h_t^n\rangle_H - \int_0^t {}_{V^*}\langle A(s,z_s^n) - A(s,z_s), h_s^n\rangle_V \mathrm{d}s \\
&- \int_0^t \langle B(s,z_s^n)\phi_s^n - B(s,z_s)\phi_s, h_s^n\rangle_H \mathrm{d}s \\
=:&I_1 + I_2 + I_3
\end{aligned}
\tag{3.12}
$$

By using the Hölder inequality, $(A3)$ and (3.7) we have

$$
\begin{aligned}
I_1 &\le \|z_t^n - z_t\|_H \cdot \|h_t^n\|_H \le \frac{1}{4}\|z_t^n - z_t\|_H^2 + \|h_t^n\|_H^2. \\
I_2 &\le \int_0^t \|A(s,z_s^n) - A(s,z_s)\|_{V^*}\|h_s^n\|_V \mathrm{d}s \\
&\le \left(\int_0^t \|A(s,z_s^n) - A(s,z_s)\|_{V^*}^{\frac{\alpha}{\alpha-1}} \mathrm{d}s\right)^{\frac{\alpha-1}{\alpha}} \left(\int_0^t \|h_s^n\|_V^\alpha \mathrm{d}s\right)^{\frac{1}{\alpha}} \\
&\le \left(\int_0^t C\left(1 + \|z_s\|_V^\alpha + \|z_s^n\|_V^\alpha\right)\mathrm{d}s\right)^{\frac{\alpha-1}{\alpha}} \left(\int_0^t \|h_s^n\|_V^\alpha \mathrm{d}s\right)^{\frac{1}{\alpha}} \\
&\le C\left(\int_0^t \|h_s^n\|_V^\alpha \mathrm{d}s\right)^{\frac{1}{\alpha}}. \qquad (3.13) \\
I_3 &\le \int_0^t \|B(s,z_s^n)\phi_s^n - B(s,z_s)\phi_s\|_H \cdot \|h_s^n\|_H \mathrm{d}s \\
&\le \sup_{s\in[0,t]} \|h_s^n\|_H \int_0^t \|B(s,z_s^n)\phi_s^n - B(s,z_s)\phi_s\|_H \mathrm{d}s \\
&\le \sup_{s\in[0,t]} \|h_s^n\|_H \left\{ N^{1/2}\left(\int_0^t \|B(s,z_s^n)\|_2^2 \mathrm{d}s\right)^{1/2} + N^{1/2}\left(\int_0^t \|B(s,z_s)\|_2^2 \mathrm{d}s\right)^{1/2}\right\} \\
&\le C \sup_{s\in[0,t]} \|h_s^n\|_H.
\end{aligned}
$$

where C is a constant which come from the following estimate

$$
\int_0^t \|B(s,z_s^n)\|_2^2 \mathrm{d}s \le C\int_0^t \left(1 + \|z_s^n\|_H^2 + \|z_s^n\|_V^\alpha\right)\mathrm{d}s \le \textbf{Constant} < \infty
$$

Combining (3.9) and (3.12)-(3.13) we have

(3.14)

$$\|z_t^n - z_t\|_H^2 + \delta \int_0^t \|z_t^n - z_t\|_V^\alpha \mathrm{d}t$$

$$\leq C \int_0^t (1 + \|\phi_s^n\|_U^2)\|z_s^n - z_s\|_H^2 \mathrm{d}s + C \left(\sup_{s\in[0,t]} \|h_s^n\|_H + \sup_{s\in[0,t]} \|h_s^n\|_H^2 + \left(\int_0^t \|h_s^n\|_V^\alpha \mathrm{d}s \right)^{\frac{1}{\alpha}} \right)$$

Then by the Gronwall lemma and L^2-boundedness of ϕ^n, there exists a constant C such that

$$\|z^n - z\| \leq C \left(\sup_{s\in[0,T]} \|h_s^n\|_H + \sup_{s\in[0,T]} \|h_s^n\|_H^2 + \left(\int_0^T \|h_s^n\|_V^\alpha \mathrm{d}s \right)^{\frac{1}{\alpha}} \right).$$

Since $h^n \to 0$ in $C([0,T];H) \cap L^\alpha([0,T];V)$, we know $z^n \to z$ strongly in $C([0,T];H) \cap L^\alpha([0,T];V)$ as $n \to \infty$.

Since Lemma 3.1 shows that the convergence of the corresponding solution z^ϕ is uniformly on S_N w.r.t. the approximation on ϕ, the conclusion on the case $\phi^n, \phi \in L^2([0,T];U)$ can de derived by the proof above and the standard 3ε-argument.

Step 2: Now we prove the conclusion for general B without assuming $(A5)$. Denote $z_{t,n}^\phi$ the solution of the following equation

$$\frac{\mathrm{d}z_{t,n}^\phi}{\mathrm{d}t} = A(t, z_{t,n}^\phi) + P_n B(t, z_{t,n}^\phi)\phi_t, \quad z_{0,n}^\phi = x,$$

where P_n is the standard projection(see $(A4)$ and Section 4 for details). By using the same argument in Lemma 3.1 we can prove

$$\begin{aligned} &\|z_n^\phi - z^\phi\|^2 + \delta \int_0^T \|z_{s,n}^\phi - z_s^\phi\|_V^\alpha \mathrm{d}s \\ &\leq \exp\left\{ \int_0^T (K + 2\|\phi_s\|_U^2)\mathrm{d}s \right\} \int_0^T \|(I - P_n)B(s, z_s^\phi)\|_2^2 \mathrm{d}s \end{aligned} \tag{3.15}$$

Since $B(\cdot,\cdot)$ are Hilbert-Schmidt (hence compact) operators, then by the dominated convergence theorem we know

$$\int_0^T \|(I - P_n)B(s, z_s^\phi)\|_2^2 \mathrm{d}s \to 0 \text{ as } n \to \infty.$$

Hence $z_n^\phi \to z^\phi$ in $C([0,T];H) \cap L^\alpha([0,T];V)$ as $n \to \infty$. Moreover, this convergence is uniformly (w.r.t ϕ) on bounded set of $L^2([0,T];U)$, which follows from (3.15) and (3.8).

Note that $P_n B$ satisfy $(A5)$, by combining with **Step 1** and standard 3ε-argument we can conclude that $z^n \to z$ strongly in $C([0,T];H) \cap L^\alpha([0,T];V)$ for general B. Hence the proof is complete. □

Lemma 3.3. *Assume (A1) − (A3) and (A5) hold. Let $\{v^\varepsilon\}_{\varepsilon>0} \subset \mathcal{A}_N$ for some $N < \infty$. Assume v^ε converge to v in distribution as S_N-valued random elements, then*

$$g^\varepsilon \left(W_\cdot + \frac{1}{\varepsilon} \int_0^\cdot v_s^\varepsilon \mathrm{d}s \right) \to g^0 \left(\int_0^\cdot v_s \mathrm{d}s \right)$$

in distribution as $\varepsilon \to 0$.

Proof. By the Girsanov theorem and uniqueness of solution to (2.3), it's easy to see that $X^\varepsilon := g^\varepsilon \left(W_\cdot + \frac{1}{\varepsilon} \int_0^\cdot v_s^\varepsilon \mathrm{d}s \right)$ (the abuse of notation here is for simplicity) is the unique solution of the following equation

$$\mathrm{d}X_t^\varepsilon = (A(t, X_t^\varepsilon) + B(t, X_t^\varepsilon) v_t^\varepsilon)\,\mathrm{d}t + \varepsilon B(t, X_t^\varepsilon) \mathrm{d}W_t,\ \ X_0^\varepsilon = x. \tag{3.16}$$

Now we only need to show $X^\varepsilon \to z^v$ in distribution as $\varepsilon \to 0$. We may assume $\varepsilon \le \frac{1}{2}$, by using the Itô formula, Young's inequality and $(A2)$ we have

$$\begin{aligned}
\mathrm{d}\|X_t^\varepsilon - z_t^v\|_H^2 =& 2{}_{V^*}\langle A(t, X_t^\varepsilon) - A(t, z_t^v), X_t^\varepsilon - z_t^v \rangle_V \mathrm{d}t \\
&+ 2\langle X_t^\varepsilon - z_t^v, (B(t, X_t^\varepsilon) - B(t, z_t^v)) v_t^\varepsilon + B(t, z_t^v)(v_t^\varepsilon - v_t) \rangle_H \mathrm{d}t \\
&+ \varepsilon^2 \|B(t, X_t^\varepsilon)\|_2^2 \mathrm{d}t + 2\varepsilon \langle X_t^\varepsilon - z_t^v, B(t, X_t^\varepsilon) \mathrm{d}W_t \rangle_H \\
\le& \left(2{}_{V^*}\langle A(t, X_t^\varepsilon) - A(t, z_t^v), X_t^\varepsilon - z_t^v \rangle_V + \|B(t, X_t^\varepsilon) - B(t, z_t^v)\|_2^2 \right) \mathrm{d}t \\
&+ 2\|v_t^\varepsilon\|_U^2 \|X_t^\varepsilon - z_t^v\|_H^2 \mathrm{d}t + 2\langle X_t^\varepsilon - z_t^v, B(t, z_t^v)(v_t^\varepsilon - v_t) \rangle_H \mathrm{d}t \\
&+ 2\varepsilon^2 \|B(t, z_t^v)\|_2^2 \mathrm{d}t + 2\varepsilon \langle X_t^\varepsilon - z_t^v, B(t, X_t^\varepsilon) \mathrm{d}W_t \rangle_H \\
\le& \left[-\delta \|X_t^\varepsilon - z_t^v\|_V^\alpha + C(1 + \|v_t^\varepsilon\|_U^2) \|X_t^\varepsilon - z_t^v\|_H^2 + 2\varepsilon^2 \|B(t, z_t^v)\|_2^2 \right] \mathrm{d}t \\
&+ 2\langle X_t^\varepsilon - z_t^v, B(t, z_t^v)(v_t^\varepsilon - v_t) \rangle_H \mathrm{d}t + 2\varepsilon \langle X_t^\varepsilon - z_t^v, B(t, X_t^\varepsilon) \mathrm{d}W_t \rangle_H.
\end{aligned} \tag{3.17}$$

Similarly we define

$$h_t^\varepsilon = \int_0^t B(s, z_s^v)(v_s^\varepsilon - v_s) \mathrm{d}s,$$

then we know that $h^\varepsilon \to 0$ in distribution as $C([0, T]; V)$-valued random element, consequently also in $C([0, T]; H) \cap L^\alpha([0, T]; V)$. Note that

$$2\langle X_t^\varepsilon - z_t^v, h_t^\varepsilon \rangle_H = \|X_t^\varepsilon - z_t^v + h_t^\varepsilon\|_H^2 - \|X_t^\varepsilon - z_t^v\|_H^2 - \|h_t^\varepsilon\|_H^2.$$

By using the Itô formula for corresponding square norm we can derive that

$$\begin{aligned}
&\int_0^t \langle X_s^\varepsilon - z_s^v, B(s, z_s^v)(v_s^\varepsilon - v_s) \rangle_H \mathrm{d}s \\
=& \langle X_t^\varepsilon - z_t^v, h_t^\varepsilon \rangle_H - \int_0^t {}_{V^*}\langle A(s, X_s^\varepsilon) - A(s, z_s^v), h_s^\varepsilon \rangle_V \mathrm{d}s \\
&- \int_0^t \langle B(s, X_s^\varepsilon) v_s^\varepsilon - B(s, z_s^v) v_s, h_s^\varepsilon \rangle_H \mathrm{d}s - \varepsilon \int_0^t \langle B(s, X_s^\varepsilon) \mathrm{d}W_s, h_s^\varepsilon \rangle_H.
\end{aligned} \tag{3.18}$$

By using the same argument as in (3.13) we obtain

$$
\begin{aligned}
&\int_0^t \langle X_s^\varepsilon - z_s^v, B(s,z_s^v)(v_s^\varepsilon - v_s)\rangle_H \mathrm{d}s\\
\le&\frac{1}{4}\|X_t^\varepsilon - z_t^v\|_H^2 + \sup_{s\in[0,t]}\|h_s^\varepsilon\|_H^2 - \varepsilon\int_0^t \langle B(s,X_s^\varepsilon)\mathrm{d}W_s, h_s^\varepsilon\rangle_H\\
&+C\left(\int_0^t (1+\|z_s^v\|_V^\alpha + \|X_s^\varepsilon\|_V^\alpha)\,\mathrm{d}s\right)^{\frac{\alpha-1}{\alpha}}\cdot\left(\int_0^t \|h_s^\varepsilon\|_V^\alpha \mathrm{d}s\right)^{\frac{1}{\alpha}}\\
&+C\sup_{s\in[0,t]}\|h_s^\varepsilon\|_H\left\{\left(\int_0^t \|B(s,X_s^\varepsilon)\|_2^2\mathrm{d}s\right)^{1/2} + \left(\int_0^t \|B(s,z_s^v)\|_2^2\mathrm{d}s\right)^{1/2}\right\}.
\end{aligned}
\tag{3.19}
$$

Hence from (3.17)-(3.19) we have

$$
\begin{aligned}
&\|X_t^\varepsilon - z_t^v\|_H^2 + \delta\int_0^t \|X_t^\varepsilon - z_t^v\|_V^\alpha \mathrm{d}s\\
&\le c_1\int_0^t (1+\|v_s^\varepsilon\|_U^2)\|X_t^\varepsilon - z_t^v\|_H^2\mathrm{d}s + c_2(\varepsilon^2 + \sup_{s\in[0,t]}\|h_s^\varepsilon\|_H^2)\\
&+c_3\left(1+\int_0^t \|X_s^\varepsilon\|_V^\alpha\mathrm{d}s\right)^{\frac{\alpha-1}{\alpha}}\cdot\left(\int_0^t \|h_s^\varepsilon\|_V^\alpha\mathrm{d}s\right)^{\frac{1}{\alpha}}\\
&+c_4\sup_{s\in[0,t]}\|h_s^\varepsilon\|_H\left\{1+\left(\int_0^t \|X_s^\varepsilon\|_H^2\mathrm{d}s\right)^{1/2}\right\}\\
&+4\varepsilon\int_0^t \langle X_s^\varepsilon - z_s^v - h_s^\varepsilon, B(s,X_s^\varepsilon)\mathrm{d}W_s\rangle_H,
\end{aligned}
\tag{3.20}
$$

where we used the estimate (see (3.6)-(3.8)) that there exists constant C such that

$$\int_0^T \|B(s,z_s^v)\|_2^2\mathrm{d}s + \int_0^T \|z_s^v\|_V^\alpha\mathrm{d}s \le C, \quad a.s.$$

By applying the Gronwall lemma we have

$$
\begin{aligned}
&\sup_{s\in[0,t]}\|X_s^\varepsilon - z_s^v\|_H^2 + \delta\int_0^t \|X_s^\varepsilon - z_s^v\|_V^\alpha\mathrm{d}s\\
&\le C\Bigg[\varepsilon^2 + \sup_{s\in[0,t]}\|h_s^\varepsilon\|_H^2 + \left(1+\int_0^t \|X_s^\varepsilon\|_V^\alpha\mathrm{d}s\right)^{\frac{\alpha-1}{\alpha}}\left(\int_0^t \|h_s^\varepsilon\|_V^\alpha\mathrm{d}s\right)^{\frac{1}{\alpha}}\\
&+\sup_{s\in[0,t]}\|h_s^\varepsilon\|_H\left\{1+\left(\int_0^t \|X_s^\varepsilon\|_H^2\mathrm{d}s\right)^{1/2}\right\} + \sup_{u\in[0,t]}\left|\varepsilon\int_0^u \langle X_s^\varepsilon - z_s^v - h_s^\varepsilon, B(s,X_s^\varepsilon)\mathrm{d}W_s\rangle_H\right|\Bigg]
\end{aligned}
\tag{3.21}
$$

Define the stopping time

$$\tau^{M,\varepsilon} = \inf\left\{t \le T : \sup_{s\in[0,t]} \|X^\varepsilon_s\|_H^2 + \int_0^t \|X^\varepsilon_s\|_V^\alpha \mathrm{d}s > M\right\}.$$

By the Burkhölder-Davis-Gundy inequality one has

$$\begin{aligned}
&\varepsilon\mathbf{E}\sup_{t\in[0,\tau^{M,\varepsilon}]}\left|\int_0^t \langle X^\varepsilon_s - z^v_s - h^\varepsilon_s, B(s,X^\varepsilon_s)\mathrm{d}W_s\rangle_H\right| \\
\le&3\varepsilon\mathbf{E}\left\{\int_0^{\tau^{M,\varepsilon}} \|X^\varepsilon_s - z^v_s - h^\varepsilon_s\|_H^2\|B(s,X^\varepsilon_s)\|_2^2\mathrm{d}s\right\}^{1/2} \\
\le&3\varepsilon\mathbf{E}\left\{\sup_{s\in[0,\tau^{M,\varepsilon}]}\|X^\varepsilon_s - z^v_s - h^\varepsilon_s\|_H^2 + C\int_0^{\tau^{M,\varepsilon}}\left(1+\|X^\varepsilon_s\|_H^2+\|X^\varepsilon_s\|_V^\alpha\right)\mathrm{d}s\right\} \\
\le&C\varepsilon \to 0\ (\varepsilon\to 0).
\end{aligned}\tag{3.22}$$

By using the similar argument in (3.17) we have

$$\mathrm{d}\|X^\varepsilon_t\|_H^2 \le -\frac{\delta}{2}\|X^\varepsilon_t\|_V^\alpha\mathrm{d}t + C(1+\|X^\varepsilon_t\|_H^2+\|v^\varepsilon_t\|_U^2\|X^\varepsilon_t\|_H^2)\mathrm{d}t + 2\varepsilon\langle X^\varepsilon_t, B(t,X^\varepsilon_t)\mathrm{d}W_t\rangle_H,$$

where C is a constant. Repeat the same argument in [21, Theorem 3.10] we can prove

$$\sup_{\varepsilon\in[0,1)}\mathbf{E}\left\{\sup_{t\in[0,T]}\|X^\varepsilon_t\|_H^2 + \int_0^T\|X^\varepsilon_t\|_V^\alpha\mathrm{d}t\right\} < \infty.$$

Hence there exists a suitable constant C such that

$$\liminf_{\varepsilon\to 0}\mathbf{P}\{\tau^{M,\varepsilon}=T\} \ge 1-\frac{C}{M}.\tag{3.23}$$

Recall that $h^\varepsilon \to 0$ in distribution in $C([0,T];H)\cap L^\alpha([0,T];V)$, combining with (3.21)-(3.23) one can conclude

$$\sup_{t\in[0,T]}\|X^\varepsilon_t - z^v_t\|_H^2 + \int_0^T\|X^\varepsilon_t - z^v_t\|_V^\alpha\mathrm{d}t \to 0\ (\varepsilon\to 0)$$

in distribution. Hence the proof is complete. □

Remark 3.1. According to Lemma 1.1, Lemma 3.2 and Lemma 3.3, we know that $\{X^\varepsilon\}$ satisfy LDP provided $(A1)-(A3)$ and $(A5)$ hold. By using some approximation argument, we can replace $(A5)$ by the weaker assumption $(A4)$.

4 Replace $(A5)$ by $(A4)$

For any fixed $n \geq 1$, let $H_n \subseteq V$ compact and $P_n : H \to H_n$ be the orthogonal projection. Let $X_t^{\varepsilon,n}$ be the solution of

$$\mathrm{d}X_t^{\varepsilon,n} = A(t, X_t^{\varepsilon,n})\mathrm{d}t + \varepsilon P_n B(t, X_t^{\varepsilon,n})\mathrm{d}W_t, \quad X_0^{\varepsilon,n} = x. \tag{4.1}$$

Since $P_n B$ satisfy $(A5)$, according to the Section 3(Remark 3.1) we know $\{X^{\varepsilon,n}\}$ satisfy the LDP provided $(A1)-(A3)$. Now we prove that $\{X^{\varepsilon,n}\}$ are the exponential good approximation to $\{X^\varepsilon\}$ if the following assumption hold.

$(A4')$

$$a_n := \sup_{(t,v)\in[0,T]\times V} \|P_n B(t,v) - B(t,v)\|_2^2 \to 0 \ (n \to \infty).$$

Lemma 4.1. *If $(A1)-(A3)$ and $(A4')$ hold, then $\forall \sigma > 0$*

$$\limsup_{n\to\infty} \limsup_{\varepsilon\to 0} \varepsilon^2 \log \mathbf{P}\left(\rho(X^\varepsilon, X^{\varepsilon,n}) > \sigma\right) = -\infty, \tag{4.2}$$

where ρ is the metric on $C([0,T];H) \cap L^\alpha([0,T];V)$ defined in (2.4).

Proof. For $\varepsilon < \frac{1}{2}$, by using the Itô formula and $(A2)$ we have

$$\begin{aligned}
&\mathrm{d}\|X_t^\varepsilon - X_t^{\varepsilon,n}\|_H^2 \\
=&\left(2_{V^*}\langle A(t,X_t^\varepsilon) - A(t,X_t^{\varepsilon,n}), X_t^\varepsilon - X_t^{\varepsilon,n}\rangle_V + \varepsilon^2\|B(t,X_t^\varepsilon) - P_n B(t,X_t^{\varepsilon,n})\|_2^2\right)\mathrm{d}t \\
&+ 2\varepsilon\langle X_t^\varepsilon - X_t^{\varepsilon,n}, (B(t,X_t^\varepsilon) - P_n B(t,X_t^{\varepsilon,n}))\mathrm{d}W_t\rangle_H,
\end{aligned}$$

where C is a constant. Define

$$\|X_t^\varepsilon - X_t^{\varepsilon,n}\| = \|X_t^\varepsilon - X_t^{\varepsilon,n}\|_H^2 + \delta\int_0^t \|X_s^\varepsilon - X_s^{\varepsilon,n}\|_V^\alpha \mathrm{d}s.$$

Note that

$$M_t^{(n)} := \int_0^t \langle X_s^\varepsilon - X_s^{\varepsilon,n}, (B(s,X_s^\varepsilon) - P_n B(s,X_s^{\varepsilon,n}))\,\mathrm{d}W_t\rangle_H$$

is a local martingale and its quadratic variation process satisfies

$$\mathrm{d}\langle M^{(n)}\rangle_t \leq 2\|X_t^\varepsilon - X_t^{\varepsilon,n}\|_H^2(\|B(t,X_t^\varepsilon) - B(t,X_t^{\varepsilon,n})\|_2^2 + a_n)\mathrm{d}t.$$

Let $\varphi_\theta(y) = (a_n + y)^\theta$ for some $\theta \leq \frac{1}{4\varepsilon^2}$, then by $(A2)$

(4.3)

$$\begin{aligned}
&\mathrm{d}\varphi_\theta(\|X_t^\varepsilon - X_t^{\varepsilon,n}\|) \\
\leq&\theta(a_n + \|X_t^\varepsilon - X_t^{\varepsilon,n}\|)^{\theta-1}\left(\mathrm{d}\|X_t^\varepsilon - X_t^{\varepsilon,n}\|_H^2 + \delta\|X_t^\varepsilon - X_t^{\varepsilon,n}\|_V^\alpha\mathrm{d}t\right) \\
&+ 4\varepsilon^2\theta(\theta-1)(a_n + \|X_t^\varepsilon - X_t^{\varepsilon,n}\|)^{\theta-2}\|X_t^\varepsilon - X_t^{\varepsilon,n}\|_H^2\left(\|B(t,X_t^\varepsilon) - B(t,X_t^{\varepsilon,n})\|_2^2 + a_n\right)\mathrm{d}t \\
\leq&C\theta\varphi_\theta\left(\|X_t^\varepsilon - X_t^{\varepsilon,n}\|\right)\mathrm{d}t + \mathrm{d}\beta_t
\end{aligned}$$

where C is a constant and β_t is a local martingale. By standard localization argument we may assume β_t is a martingale for simplicity. Let $\theta = \frac{1}{4\varepsilon^2}$ we know

$$N_t := \exp\left[-\frac{C}{4\varepsilon^2}t\right]\varphi_{\frac{1}{4\varepsilon^2}}\left(\|X_t^\varepsilon - X_t^{\varepsilon,n}\|\right)$$

is a supermartingale. Hence we have

$$\begin{aligned}
&\mathbf{P}\left(\rho(X^\varepsilon, X^{\varepsilon,n}) > 2\sigma\right)\\
\leq&\mathbf{P}\left(\sup_{t\in[0,T]}\|X_t^\varepsilon - X_t^{\varepsilon,n}\|_H > \sigma\right) + \mathbf{P}\left(\int_0^T \|X_t^\varepsilon - X_t^{\varepsilon,n}\|_V^\alpha \mathrm{d}t > \sigma^\alpha\right)\\
\leq&\mathbf{P}\left(\sup_{t\in[0,T]} N_t > \exp\left[-\frac{C}{4\varepsilon^2}T\right](\sigma^2 + a_n)^{\frac{1}{4\varepsilon^2}}\right) + \mathbf{P}\left(\sup_{t\in[0,T]} N_t > \exp\left[-\frac{C}{4\varepsilon^2}T\right](\delta\sigma^\alpha + a_n)^{\frac{1}{4\varepsilon^2}}\right)\\
\leq&\exp\left[\frac{C}{4\varepsilon^2}T\right](\sigma^2 + a_n)^{-\frac{1}{4\varepsilon^2}}\mathbf{E}N_0 + \exp\left[\frac{C}{4\varepsilon^2}T\right](\delta\sigma^\alpha + a_n)^{-\frac{1}{4\varepsilon^2}}\mathbf{E}N_0\\
=&\exp\left[\frac{C}{4\varepsilon^2}T\right]\left\{\left(\frac{a_n}{\sigma^2 + a_n}\right)^{\frac{1}{4\varepsilon^2}} + \left(\frac{a_n}{\delta\sigma^\alpha + a_n}\right)^{\frac{1}{4\varepsilon^2}}\right\}.
\end{aligned}$$

This implies that

$$\begin{aligned}
&\limsup_{\varepsilon\to 0}\varepsilon^2 \log \mathbf{P}\left(\rho(X^\varepsilon, X^{\varepsilon,n}) > 2\sigma\right)\\
&\leq \frac{CT}{4} + \max\left\{\log\frac{a_n}{\sigma^2 + a_n}, \log\frac{a_n}{\delta\sigma^\alpha + a_n}\right\}.
\end{aligned}$$

Since $(A4')$ says $a_n \to 0$ as $n \to \infty$, (4.2) hold and the proof is complete.

□

Corollary 4.2. *If* $(A1)-(A3)$ *and* $(A4')$ *hold, then* $\{X^\varepsilon\}$ *satisfy the LDP in* $C([0,T];H) \cap L^\alpha([0,T];V)$ *with rate function (2.7).*

Proof. According to [40, Theorem 2.1] and section 3 one can conclude $\{X^\varepsilon\}$ satisfy the LDP with the following rate function

$$\tilde{I}(f) := \sup_{r>0}\liminf_{n\to\infty}\inf_{g\in S_r(f)} I^n(g) = \sup_{r>0}\limsup_{n\to\infty}\inf_{g\in S_r(f)} I^n(g).$$

where $S_r(f)$ is the closed ball in $C([0,T];H) \cap L^\alpha([0,T];V)$ centered at f with radius r and I^n is given by

$$I^n(z) := \inf\left\{\frac{1}{2}\int_0^T \|\phi_s\|_U^2 \mathrm{d}s : \ z = z^{n,\phi},\ \phi \in L^2([0,T],U)\right\}, \tag{4.4}$$

where $z^{n,\phi}$ is the unique solution of following equation

$$\frac{\mathrm{d}z_t^n}{\mathrm{d}t} = A(t, z_t^n) + P_n B(t, z_t^n)\phi_t,\ \ z_0^n = x.$$

Now we only need to prove $\tilde{I} = I$, i.e.

$$I(f) = \sup_{r>0} \liminf_{n\to\infty} \inf_{g\in S_r(f)} I^n(g).$$

We will first show that for any $r > 0$

$$I(f) \geq \liminf_{n\to\infty} \inf_{g\in S_r(f)} I^n(g).$$

We assume $I(f) < \infty$, then by Lemma 3.2 there exists ϕ such that

$$f = z^{\phi} \ \text{ and } \ I(f) = \frac{1}{2}\int_0^T \|\phi_s\|_U^2 \mathrm{d}s.$$

Since $z^{n,\phi} \to z^{\phi}$, for n large enough we have

$$f_n := z^{n,\phi} \in S_r(f).$$

Notice $I^n(f_n) \leq \frac{1}{2}\int_0^T \|\phi_s\|_U^2 \mathrm{d}s$, hence we have

$$\liminf_{n\to\infty} \inf_{g\in S_r(f)} I^n(g) \leq \liminf_{n\to\infty} I^n(f_n) \leq I(f).$$

Since r is arbitrary we have proved the lower bound

$$I(f) \geq \sup_{r>0} \liminf_{n\to\infty} \inf_{g\in S_r(f)} I^n(g).$$

For the upper bound we can proceed as in finite dimensional case in [34, Lemma 4.6] to show

$$\limsup_{n\to\infty} \inf_{g\in S_r(f)} I^n(g) \geq \inf_{g\in S_r(f)} I(g)$$

Hence we have

$$\sup_{r>0} \limsup_{n\to\infty} \inf_{g\in S_r(f)} I^n(g) \geq \sup_{r>0} \inf_{g\in S_r(f)} I(g) \geq I(f).$$

Hence the proof is complete. □

In order to replace the assumption $(A4')$ by $(A4)$, we need to use some truncation techniques (cf. [34, 9]).

Lemma 4.3. *Assume* $(A1)-(A4)$ *hold, then*

$$\lim_{R\to\infty} \limsup_{\varepsilon\to 0} \varepsilon^2 \log \mathbf{P}\Big(\sup_{t\in[0,T]} \|X_t^{\varepsilon}\|_H^2 + \frac{\delta}{2}\int_0^T \|X_t^{\varepsilon}\|_V^{\alpha} \mathrm{d}t > R\Big) = -\infty. \tag{4.5}$$

Proof. By using the Itô formula we have

$$\mathrm{d}\|X_t^\varepsilon\|_H^2 = \big(2_{V^*}\langle A(t,X_t^\varepsilon), X_t^\varepsilon\rangle_V + \varepsilon^2\|B(t,X_t^\varepsilon)\|_2^2\big)\,\mathrm{d}t + 2\varepsilon\langle X_t^\varepsilon, (B(t,X_t^\varepsilon)\mathrm{d}W_t\rangle_H.$$

Note that $M_t^{(n)} := \int_0^t \langle X_s^\varepsilon, B(s,X_s^\varepsilon)\mathrm{d}W_s\rangle_H$ is a local martingale and

$$d\langle M^{(n)}\rangle_t \le \|X_t^\varepsilon\|_H^2\|B(t,X_t^\varepsilon)\|_2^2\mathrm{d}t.$$

Define

$$\|X_t^\varepsilon\| := \|X_t^\varepsilon\|_H^2 + \frac{\delta}{2}\int_0^t \|X_s^\varepsilon\|_V^\alpha \mathrm{d}s, \quad \varphi_\theta(y) = (1+y)^\theta, \ \theta > 0,$$

then for $\theta \le \frac{1}{2\varepsilon^2}$ by $(A2)$ and $(A3)$ we have

$$\begin{aligned}\mathrm{d}\varphi_\theta(\|X_t^\varepsilon\|) \le& \theta(1+\|X_t^\varepsilon\|)^{\theta-1}\left(\mathrm{d}\|X_t^\varepsilon\|_H^2 + \frac{\delta}{2}\|X_t^\varepsilon\|_V^\alpha \mathrm{d}t\right)\\ &+ 2\varepsilon^2\theta(\theta-1)(1+\|X_t^\varepsilon\|)^{\theta-2}\|X_t^\varepsilon\|_H^2\|B(t,X_t^\varepsilon)\|_2^2\mathrm{d}t\\ \le& C\theta\varphi_\theta(\|X_t^\varepsilon\|)\mathrm{d}t + \mathrm{d}\beta_t\end{aligned} \tag{4.6}$$

where β_t is a local martingale. We also omit the standard localization procedure here. Let $\theta = \frac{1}{2\varepsilon^2}$ we know

$$N_t := \exp\left[-\frac{C}{2\varepsilon^2}t\right]\varphi_{\frac{1}{2\varepsilon^2}}(\|X_t^\varepsilon\|)$$

is a supermartingale. Hence we have

$$\begin{aligned}&\mathbf{P}\left(\sup_{t\in[0,T]}\|X_t^\varepsilon\|_H^2 + \frac{\delta}{2}\int_0^T \|X_t^\varepsilon\|_V^\alpha \mathrm{d}t > R\right)\\ \le&\mathbf{P}\left(\sup_{t\in[0,T]} N_t > \exp\left[-\frac{C}{2\varepsilon^2}T\right](1+R)^{\frac{1}{2\varepsilon^2}}\right)\\ \le&\exp\left[\frac{C}{2\varepsilon^2}T\right](1+R)^{-\frac{1}{2\varepsilon^2}}\mathbf{E}N_0\\ =&\exp\left[\frac{C}{2\varepsilon^2}T\right]\left(\frac{1}{1+R}\right)^{\frac{1}{2\varepsilon^2}}.\end{aligned}$$

This implies that

$$\limsup_{\varepsilon\to 0}\varepsilon^2\log\mathbf{P}\left(\sup_{t\in[0,T]}\|X_t^\varepsilon\| > R\right) \le \frac{1}{2}\log\frac{1}{1+R} + \frac{CT}{2}.$$

Therefore, (4.5) hold. □

After all these preparations, now we can finish the proof of Theorem 2.1.

Proof of Theorem 2.1: The proof here is a slight modification of [34, Theorem 4.13]. Define $\xi : V \to [0,1]$ be a C_0^∞-function such that

$$\xi(v) := \begin{cases} 0, & \text{if } \|v\|_H > 2, \\ 1, & \text{if } \|v\|_H \le 1. \end{cases}$$

Let $\xi_N(v) = \xi(\frac{v}{N})$ and

$$B_N(t,v) = \xi_N(v)B(t,v) + (1-\xi_N(v))B(t,0).$$

Consider the mollified problem for equation (2.3):

$$\mathrm{d}X_{t,N}^\varepsilon = A(t, X_{t,N}^\varepsilon)\mathrm{d}t + \varepsilon B_N(t, X_{t,N}^\varepsilon)\mathrm{d}W_t, \ X_0 = x. \tag{4.7}$$

It's easily to see that A, B_N satisfy $(A1)-(A3)$ and $(A4')$, since in this case $(A4)$ implies that for B_N

$$a_n = \max\left\{ \sup_{(t,v)\in[0,T]\times S_{2N}} \|(I-P_n)B(t,v)\|_2^2, \ \sup_{t\in[0,T]} \|(I-P_n)B(t,0)\|_2^2 \right\} \to 0 (n\to\infty).$$

Hence by Corollary 4.2 we know $\{X_N^\varepsilon\}_{\varepsilon>0}$ satisfy large deviation principle on $C([0,T];H)\cap L^\alpha([0,T];V)$ with the following mollified rate function

$$I_N(z) := \inf\left\{ \frac{1}{2}\int_0^T \|\phi_s\|_U^2 \mathrm{d}s : \ z = z_N^\phi, \ \phi \in L^2([0,T],U) \right\}, \tag{4.8}$$

where z_N^ϕ is the unique solution of following equation

$$\frac{\mathrm{d}z_{t,N}}{\mathrm{d}t} = A(t, z_{t,N}) + B_N(t, z_{t,N})\phi_t, \ z_{0,N} = x.$$

Let $N \to \infty$, then the LDP for $\{X^\varepsilon\}$ can be derived as in the finite dimensional case.

According to Lemma 3.2, I defined in (2.7) is a (good) rate function. Note $I_N(z) = I(z)$ for any $z \in C([0,T];H) \cap L^\alpha([0,T];V)$ satisfy

$$\|z\|_T := \sup_{t\in[0,T]} \|z_t\|_H \le N.$$

We now first show that for any open set $G \subseteq C([0,T];H) \cap L^\alpha([0,T];V)$

$$\liminf_{\varepsilon\to0} \varepsilon^2 \log \mathbf{P}\left(X^\varepsilon \in G\right) \ge -\inf_{z\in G} I(z).$$

Obviously, we only need to prove that for all $\overline{z} \in G$ with $\overline{z}_0 = x$

$$\liminf_{\varepsilon\to0} \varepsilon^2 \log \mathbf{P}\left(X^\varepsilon \in G\right) \ge -I(\overline{z}).$$

Choose $R > 0$ such that $\|\overline{z}\|_T < R$ and set

$$N_R = \{z \in C([0,T]; H) \cap L^\alpha([0,T]; V) : \|z\|_T < R\}.$$

Then we have

$$\begin{aligned}
\liminf_{\varepsilon\to 0} \varepsilon^2 \log \mathbf{P}\left(X^\varepsilon \in G\right) &\geq \liminf_{\varepsilon\to 0} \varepsilon^2 \log \mathbf{P}\left(X^\varepsilon \in G \cap N_R\right) \\
&= \liminf_{\varepsilon\to 0} \varepsilon^2 \log \mathbf{P}\left(X_N^\varepsilon \in G \cap N_R\right) \\
&\geq - \inf_{z\in G\cap N_R} I_N(z) \\
&\geq -I(\overline{z}).
\end{aligned}$$

Finally, given a closed set F and an $L < \infty$, by Lemma 4.3 there exists R such that

$$\begin{aligned}
\limsup_{\varepsilon\to 0} \varepsilon^2 \log \mathbf{P}\left(X^\varepsilon \in F\right) &\leq \limsup_{\varepsilon\to 0} \varepsilon^2 \log \left(\mathbf{P}(X^\varepsilon \in F \cap \overline{N_R}) + \mathbf{P}(X^\varepsilon \in N_R^c)\right) \\
&\leq (- \inf_{z\in F\cap\overline{N_R}} I_N(z)) \vee (-L) \\
&\leq - \left[\inf_{z\in F} I(z) \wedge L\right].
\end{aligned}$$

Let $L \to \infty$, we obtain

$$\limsup_{\varepsilon\to 0} \varepsilon^2 \log \mathbf{P}\left(X^\varepsilon \in F\right) \leq - \inf_{z\in F} I(z).$$

Now the proof of Theorem 2.1 is complete. □

5 Examples

Now we can apply the main results to many stochastic evolution equations as applications. As a preparation we prove the following lemma first.

Lemma 5.1. *Let $(E, \langle\cdot,\cdot\rangle, \|\cdot\|)$ is a Hilbert space, then for any $r \geq 0$ we have*

$$\langle \|a\|^r a - \|b\|^r b, a - b\rangle \geq 2^{-r}\|a - b\|^{r+2}, \;\; a, b \in E. \tag{5.1}$$

$$\|\|a\|^{r-1}a - \|b\|^{r-1}b\| \leq \max\{r, 1\}\|a - b\|(\|a\|^{r-1} + \|b\|^{r-1}), \;\; a, b \in E. \tag{5.2}$$

If $0 < r < 1$, then there exists a constant $C > 0$ such that

$$||a|^{r-1}a - |b|^{r-1}b| \leq C|a - b|^r, \;\; a, b \in \mathbb{R}. \tag{5.3}$$

Proof. (i) By the symmetry of (5.1) we may assume $\|a\| \geq \|b\|$. Then

$$\begin{aligned}
& \langle \|a\|^r a - \|b\|^r b, a - b\rangle \\
= & \|b\|^r \|a-b\|^2 + (\|a\|^r - \|b\|^r)\langle a, a-b\rangle \\
= & \|b\|^r \|a-b\|^2 + (\|a\|^r - \|b\|^r) \cdot \frac{1}{2}(\|a\|^2 + \|a-b\|^2 - \|b\|^2) \\
\geq & \|b\|^r \|a-b\|^2 + \frac{1}{2}(\|a\|^r - \|b\|^r)\|a-b\|^2 \\
= & \frac{1}{2}(\|a\|^r + \|b\|^r)\|a-b\|^2 \\
\geq & 2^{-r}\|a-b\|^{r+2},
\end{aligned}$$

since $\|a-b\|^r \leq 2^{r-1}(\|a\|^r + \|b\|^r)$.

(ii) The proof of (5.2) and (5.3) is similar. $\square$

The first example is to obtain the LDP for a class of reaction-diffusion type SPDEs within the variational framework, which improve the main result in [9].

Example 5.2. *(Stochastic reaction-diffusion equations)*
Let Λ be an open bounded domain in $\mathbb{R}^d$ with smooth boundary and L be a negative definite self-adjoint operator on $H := L^2(\Lambda)$. Suppose

$$V := \mathscr{D}(\sqrt{-L}), \quad \|v\|_V := \|\sqrt{-L}v\|_H.$$

is a Banach space such that $V \subseteq H$ is dense and compact, and L can be extended as a continuous operator from V to it's dual space V^*. Consider the following semilinear stochastic equation

$$\mathrm{d}X_t^\varepsilon = (LX_t^\varepsilon + F(t, X_t^\varepsilon))\mathrm{d}t + \varepsilon B(t, X_t^\varepsilon)\mathrm{d}W_t, \ X_0^\varepsilon = x \in H, \tag{5.4}$$

where W_t is a cylindrical Wiener process on another separable Hilbert space U and

$$F : [0,T] \times V \to V^*, \quad B : [0,T] \times V \to L_2(U; V).$$

If F and B satisfy the following conditions:

$$\begin{aligned}
& 2_{V^*}\langle F(t,u) - F(t,v), u - v\rangle_V + \|B(t,u) - B(t,v)\|_2^2 \leq C\|u-v\|_H^2, \\
& \|F(t,v)\|_{V^*} \leq C(1 + \|v\|_V), \ \ \|B(t,v)\|_2 \leq C(1 + \|v\|_H^\gamma), \ u, v \in V.
\end{aligned} \tag{5.5}$$

where $C, \gamma > 0$ are constants, then $\{X^\varepsilon\}$ satisfy the large deviation principle on $C([0,T]; H) \cap L^2([0,T]; V)$.

Proof. From the assumptions (5.5), it's easy to show that $(A1) - (A4)$ hold for $\alpha = 2$. Hence the conclusion follows from Theorem 2.1. $\square$

Remark 5.1. (i) We can simply take L as the Laplace operator with Dirichlet boundary condition and $F(t,X_t) = -|X_t|^{p-2}X_t(1 \le p \le 2)$ as a concrete example.

(ii) Compare with the result in [9, Theorem 4.2](only time homogeneous case), the author in [9] need to assume F is local Lipschitz and have more restricted range conditions:

$$F : [0,T] \times V \to H.$$

In our example we can allow F is monotone and take values in V^*. Another difference is we also drop the non-degenerated condition $(A.4)$ on B in [9].

(iii) Note here one can also take $B : V \to L_2(U;H)$ with locally compact range, which seems not allowed in [9, Theorem 4.2].

The second example is stochastic porous media equations, which have been studied intensively in recent years, see e.g.[11, 27, 32, 39]. We use the same framework as in [32, 39].

Example 5.3. *(Stochastic porous media equations)*
Let $(E, \mathcal{M}, \mathbf{m})$ be a separable probability space and $(L, \mathscr{D}(L))$ a negative definite self-adjoint linear operator on $(L^2(\mathbf{m}), \langle\cdot,\cdot\rangle)$ with spectrum contained in $(-\infty, -\lambda_0]$ for some $\lambda_0 > 0$. Then the embedding

$$H^1 := \mathscr{D}(\sqrt{-L}) \subseteq L^2(\mathbf{m})$$

is dense and continuous. Define H is the dual Hilbert space of H^1 realized through this embedding. Assume L^{-1} is continuous on $L^{r+1}(\mathbf{m})$.

For fixed $r > 1$, we consider the following Gelfand triple

$$V := L^{r+1}(\mathbf{m}) \subseteq H \subseteq V^*$$

and the stochastic porous media equation

$$\mathrm{d}X_t^\varepsilon = (L\Psi(t,X_t^\varepsilon) + \Phi(t,X_t^\varepsilon))\mathrm{d}t + \varepsilon B(t,X_t^\varepsilon)\mathrm{d}W_t, \ \ X_0^\varepsilon = x \in H. \tag{5.6}$$

where W_t is a cylindrical Wiener process on $L^2(\mathbf{m})$, $\Psi, \Phi : [0,T] \times \mathbb{R} \to \mathbb{R}$ are measurable and continuous in the second variable. Suppose $L^2(\mathbf{m}) \subseteq H$ is compact and $B : [0,T] \times V \to L_2(L^2(\mathbf{m}))$. If there exist two constants $\delta > 0$ and K such that

$$\begin{aligned}
&|\Psi(t,x)| + |\Phi(t,x)| + \|B(t,0)\|_2 \le K(1+|x|^r), \quad t \in [0,T], x \in \mathbb{R};\\
&-\langle \Psi(t,u) - \Psi(t,v), u-v\rangle - \langle \Phi(t,u) - \Phi(t,v), L^{-1}(u-v)\rangle\\
&\qquad \le -\delta \|u-v\|_V^{r+1} + K\|u-v\|_H^2;\\
&\|B(t,u) - B(t,v)\|_2^2 \le K\|u-v\|_H^2, \quad t \in [0,T], u,v \in V.
\end{aligned} \tag{5.7}$$

Then $\{X^\varepsilon\}$ satisfy the large deviation principle on $C([0,T];H) \cap L^{r+1}([0,T];V)$.

Proof. From the assumptions and the relation

$$ {}_{V^*}\langle L\Phi(t,u)+\Phi(t,u),u\rangle_V = -\langle \Phi(t,u),u\rangle - \langle \Phi(t,u),L^{-1}u\rangle, $$

it's easy to show that $(A1)-(A4)$ hold for $\alpha = r+1$ from (5.7). We refer to [30, Example 4.1.11] for details, see also [11, 32, 39]. Hence the conclusion follows from Theorem 2.1. $\square$

Remark 5.2. (i) If we take L the Laplace operator on a smooth bounded domain in a complete Riemannian manifold with Dirichlet boundary condition. A simple example for Ψ and Φ satisfy (5.7) is given by

$$ \Psi(t,x) = f(t)|x|^{r-1}x, \quad \Phi(t,x) = g(t)x $$

for some strictly positive continuous function f and bounded function g on $[0,T]$.

(ii) This example generalized the main result in [32, Theorem 1.1] where LDP was obtained for stochastic porous media equations with additive noise. In [32] the authors mainly used the piecewise linear approximation to the path of Wiener process and generalized contraction principle.

If we assume $0 < r < 1$ in the above example (cf.[23, 27]), then the equation is the stochastic version of classical fast diffusion equation. The behavior of the solutions to these two types of PDE has many essentially different aspects, see e.g.[1].

Example 5.4. *(Stochastic fast diffusion equations)*
Assume the same framework as Example 5.3 for $0 < r < 1$, i.e. assume the embedding $V := L^{r+1}(\mathbf{m}) \subseteq H$ is continuous and dense. We consider the equation

$$ \mathrm{d}X_t^\varepsilon = \left\{ L\Psi(t,X_t^\varepsilon) + \eta_t X_t^\varepsilon \right\}\mathrm{d}t + \varepsilon B(t,X_t^\varepsilon)\mathrm{d}W_t, \ X_0^\varepsilon = x \in H, \tag{5.8} $$

where $\eta : [0,T] \to \mathbb{R}$ is locally bounded and measurable and $\Psi : [0,T]\times\mathbb{R} \to \mathbb{R}$ is measurable and continuous in the second variable, W_t is a cylindrical Wiener process on $L^2(\mathbf{m})$ and $B : [0,T]\times V \to L_2(L^2(\mathbf{m}))$ are measurable.

Suppose there exist constants $\delta > 0$ and K such that for all $x,y \in \mathbb{R}, t \in [0,T]$ and $u,v \in V$

$$ \begin{aligned} &|\Psi(t,x)| + \|B(t,0)\|_2 \le K(1+|x|^r); \\ &(\Psi(t,x)-\Psi(t,y))(x-y) \ge \delta|x-y|^2(|x|\vee|y|)^{r-1}; \\ &\|B(t,u)-B(t,v)\|_2^2 \le K\|u-v\|_H^2; \\ &\|B(t,u)\|_{L(L^2(\mathbf{m}),V^*)} \le K(1+\|u\|_V^r). \end{aligned} \tag{5.9} $$

Then $\{X^\varepsilon\}$ satisfy the large deviation principle on $C([0,T];H)$.

Proof. Note that

$$ {}_{V^*}\langle L\Psi(t,u)+\eta_t u,u\rangle_V = -\langle \Psi(t,u),u\rangle_{L^2} + \langle \eta_t u,u\rangle_H, $$

then it's easy to show $(A1),(A2'),(A3)-(A4)$ hold for $\alpha = r+1$ under assumptions (5.9). Then the conclusion follows from Theorem 2.2. $\square$

Remark 5.3. (i) In particular, if $\eta = 0, B = 0$ and $\Psi(t,s) = |s|^{r-1}s$ for some $r \in (0,1)$, then (5.8) reduces back to the classical fast-diffusion equations (cf. [1]).

(ii) In the example we assume the embedding $L^{r+1}(\mathbf{m}) \subseteq H$ is continuous and dense only for simplicity, see [23] and [30, Remark 4.1.15] for some sufficient conditions of this assumption. But in general $L^{r+1}(\mathbf{m})$ and H are incomparable, hence one need to use the more general framework as in [27] involving with Orlicz space.

Example 5.5. *(Stochastic p-Laplace equation)*
Let Λ be an open bounded domain in $\mathbb{R}^d$ with smooth boundary. We consider the triple

$$V := H_0^{1,p}(\Lambda) \subseteq H := L^2(\Lambda) \subseteq (H_0^{1,p}(\Lambda))^*$$

and the stochastic p-Laplace equation

$$\mathrm{d}X_t^\varepsilon = \left[\mathbf{div}(|\nabla X_t^\varepsilon|^{p-2}\nabla X_t^\varepsilon) - \eta_t |X_t^\varepsilon|^{\tilde{p}-2}X_t^\varepsilon\right]\mathrm{d}t + \varepsilon B(t, X_t^\varepsilon)\mathrm{d}W_t, X_0^\varepsilon = x \in H, \tag{5.10}$$

where $2 \le p < \infty, 1 \le \tilde{p} \le p$, η is positive continuous function and W_t is a cylindrical Wiener process on H. If

$$B(t,v) = \sum_{i=1}^N b_i(v)B_i(t),$$

where $b_i(\cdot) : V \to \mathbb{R}$ are Lipschitz functions and $B_i(\cdot) : [0,T] \to L_2(H)$ are continuous, then $\{X^\varepsilon\}$ satisfy the large deviation principle on $C([0,T];H) \cap L^p([0,T];V)$.

Proof. The assumptions for existence and uniqueness of the solution were verified in [30, Example 4.1.9] for $\alpha = p$. Hence we only need to prove $(A2)$ holds here. By using (5.1) in Lemma 5.1 we have

$$\begin{aligned}
& {}_{V^*}\langle \mathbf{div}(|\nabla u|^{p-2}\nabla u) - \mathbf{div}(|\nabla v|^{p-2}\nabla v), u - v\rangle_V \\
= & -\int_\Lambda \langle |\nabla u(x)|^{p-2}\nabla u(x) - |\nabla v(x)|^{p-2}\nabla v(x), \nabla u(x) - \nabla v(x)\rangle_{\mathbb{R}^d}\mathrm{d}x \\
\le & -2^{p-2}\int_\Lambda |\nabla u(x) - \nabla v(x)|^p \mathrm{d}x \\
\le & -c\|u - v\|_V^p.
\end{aligned}$$

where c is a positive constant and follows from the Poincaré inequality.

By the monotonicity of function $|x|^{\tilde{p}-2}x$ we know

$${}_{V^*}\langle |u|^{\tilde{p}-2}u - |v|^{\tilde{p}-2}v, u - v\rangle_V \ge 0.$$

Hence $(A2)$ holds. Then the conclusion follows from Theorem 2.1. □

Remark 5.4. If $1 < p < 2$ in (5.10), then the assumption $(A2)$ does not hold. Hence like the case of stochastic fast diffusion equations, we should apply Theorem 2.2 to derive the LDP for (5.10) on $C([0,T];H)$.

The following SPDE was studied in [21, 22]. The main part of drift is a high order generalization of the Laplace operator.

Example 5.6. *Let Λ is an open bounded domain in $\mathbb{R}^1$ and $m \in \mathbb{N}_+$, consider the triple*

$$V := H_0^{m,p}(\Lambda) \subseteq H := L^2(\Lambda) \subseteq (H_0^{m,p}(\Lambda))^*$$

and the stochastic evolution equation

$$\begin{aligned} \mathrm{d}X_t^\varepsilon(x) = &\left[(-1)^{m+1}\frac{\partial}{\partial x^m}\left(\left|\frac{\partial^m}{\partial x^m}X_t^\varepsilon(x)\right|^{p-2}\frac{\partial^m}{\partial x^m}X_t^\varepsilon(x)\right) + F(t, X_t(x))\right]\mathrm{d}t \\ &+ \varepsilon B(t, X_t^\varepsilon(x))\mathrm{d}W_t, \quad X_0^\varepsilon = x \in H, \end{aligned} \tag{5.11}$$

where $2 \le p < \infty$, W_t is a cylindrical Wiener process on H and

$$F : [0,T] \times V \to V^*, \quad B : [0,T] \times V \to L_2(H)$$

are measurable. Suppose $B(t,v) = QB_0(t,v), Q \in L_2(H)$ and

$$\begin{aligned} 2_{V^*}\langle F(t,u) - F(t,v), u - v\rangle_V &\le C\|u-v\|_H^2, \\ \|B_0(t,u) - B_0(t,v)\|_{L(H)} &\le C\|u-v\|_H, \\ \|F(t,u)\|_{V^*} + \|B_0(t,0)\|_{L(H)} &\le C(1 + \|u\|_V^{p-1}), \ u, v \in V, \ t \in [0,T]. \end{aligned}$$

where C is a constant. Then $\{X^\varepsilon\}$ satisfy the large deviation principle on $C([0,T]; H) \cap L^p([0,T]; V)$.

Proof. By using Lemma 5.1, $(A2)$ can be verified by the same argument as in Example 5.5. And $(A1), (A3), (A4)$ follow from the assumptions obviously, hence the conclusion follows from Theorem 2.1.

□

Acknowledgements

The author would like to thank Professor Michael Röckner, Fengyu Wang and Xicheng Zhang for their valuable discussions, and also thank Professor Paul Dupuis and Amarjit Budhiraja for the stimulating communications on their works. Many helpful comments from the referee are also gratefully acknowledged.

References

[1] D.G. Aronson, *The porous medium equation,* Lecture Notes in Mathematics 1224, Springer, Berlin, 1–46, 1986.

[2] R.G. Azencott, *Grandes deviations et applications, Ecole d'Eté de Probabilités de Saint-Flour VII,* Lecture Notes in Mathematics 774, 1980.

[3] A. Bensoussan, *Filtrage optimale des systemes linéaires,* Dunod, Paris, 1971.

[4] A. Bensoussan and R. Temam, *Equations aux derives partielles stochastiques non linéaires,* Isr. J. Math. 11(1972), 95–129.

[5] A. Budhiraja and P. Dupuis, *A variational representation for positive functionals of infinite dimensional Brownian motion,* Probab. Math. Statist. 20(2000), 39–61.

[6] A. Budhiraja, P. Dupuis and V. Maroulas, *Large deviations for infinite dimensional stochastic dynamical systems,* Ann. Probab. 36(2008), 1390-1420.

[7] W. Bryc, *Large deviations by the asymptotic value method,* In M. Pinsky, editor, Diffusion Processes and Related Problems in Analysis, vol.1, 447–472. Birkhäuser, Boston, 1990.

[8] S. Cerrai and M. Röckner, *Large deviations for stochastic reaction-diffusion systems with multiplicative noise and non-Lipschitz reaction term,* Ann. Probab. 32(2004), 1100–1139.

[9] P.L. Chow, *Large deviation problem for some parabolic Itô equations,* Commun. Pure Appl. Math. 45(1992), 97–120.

[10] G. Da Prato and M. Röckner, *Weak solutions to stochastic porous media equations,* J. Evolution Equ. 4(2004), 249–271.

[11] G. Da Prato, M. Röckner, Rozovskii and F.-Y. Wang, *Strong solutions to stochastic generalized porous media equations: existence, uniqueness and ergodicity,* Comm. Part. Diff. Equat. 31(2006), no.2, 277-291.

[12] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions,* Encyclopedia of Mathematics and its Applications, Cambridge University Press. 1992.

[13] A. Dembo and O. Zeitouni, *Large deviations techniques and applications,* Springer-Verlag, New York. 2000.

[14] M.D. Donsker and S.R.S. Varadhan, *Asymptotic evalution of certain Markov process expectations for large time, I, II, III,* Comm. Pure Appl. Math. 28(1975), 1–47; 28(1975), 279–301; 29(1977), 389–461.

[15] J. Duan and A. Millet, *Large deviations for the Boussinesq Equations under Random Influences,* Preprint.

[16] P. Dupuis and R. Ellis, *A weak convergence approach to the theory of large deviations,* Wiley, New York. 1997.

[17] J. Feng and T.G. Kurtz, *Large Deviations of Stochastic Processes,* vol. 131 of *Mathematical Surveys and Monographs,* American Mathematical Society, Providence, RI, 2006.

[18] M.I. Freidlin, *Random perturbations of reaction-diffusion equations: the quasi-deterministic approximations,* Trans. Amer. Math. Soc. 305 (1988), 665–697.

[19] M.I. Freidlin and A.D. Wentzell, *Random perturbations of dynamical systems,* Translated from the Russian by Joseph Szu"cs. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 260. Springer-Verlag, New York, 1984.

[20] I. Gyöngy and A. Millet, *On discretization schemes for stochastic evolution equations,* Pot. Anal. 23(2005), 99–134.

[21] N.V. Krylov and B.L. Rozovskii, *Stochastic evolution equations,* Translated from Itogi Naukii Tekhniki, Seriya Sovremennye Problemy Matematiki 14(1979), 71–146, Plenum Publishing Corp. 1981.

[22] W. Liu, *Harnack inequality and applications for stochastic evolution equations with monotone drifts,* SFB-Preprint 09-023.

[23] W. Liu and F.-Y. Wang, *Harnack inequality and strong Feller property for stochastic fast diffusion equations,* J. Math. Anal. Appl. 342(2008), 651-662.

[24] E. Pardoux, *Equations aux dérivées partielles stochastiques non linéaires monotones,* Thesis, Université Paris XI, 1975.

[25] S. Peszat, *Large deviation principle for stochastic evolution equations,* Probab. Theory Relat. Fields. 98(1994), 113–136.

[26] A.A. Pukhalskii, *On the theory of large deviations,* Theory probab. Appl. 38(1993), 490–497.

[27] J. Ren, M. Röckner and F.-Y. Wang, *Stochastic generalized porous media and fast diffusion equations,* J. Diff. Equat. 238(2007), 118–152.

[28] J. Ren and X. Zhang, *Freidlin-Wentzell large deviations for homeomorphism flows of non-Lipschitz SDE,* Bull. Sci. 129(2005), 643–655.

[29] J. Ren and X. Zhang, *Schilder theorem for the Brownian motion on the diffeomorphism group of the circle,* J. Funct. Anal. 224(2005), 107–133.

[30] C. Prévôt and M. Röckner, *A Concise Course on Stochastic Partial Differential Equations,* Lecture Notes in Mathematics 1905, Springer, 2007.

[31] M. Röckner, B. Schmuland and X. Zhang, *Yamada-Watanabe Theorem for stochastic evolution equations in infinite dimensions,* Condensed Matter Physics 54(2008), 247-259.

[32] M. Röckner, F.-Y. Wang and L. Wu, *Large deviations for stochastic Generalized Porous Media Equations* , Stoch. Proc. Appl. 116(2006), 1677–1689.

[33] S.S. Sritharan and P. Sundar, *Large deviations for the two-dimensional Navier-Stokes equations with multiplicative noise,* Stoc. Proc. Appl. 116(2006), 1636–1659.

[34] D.W. Stroock, *An Introduction to the Theory of Large Deviations,* Spring-Verlag, New York, 1984.

[35] S.R.S. Varadhan, *Asymptotic probabilities and differential equations,* Comm. Pure Appl. Math. 19 (1966), 261–286.

[36] S.R.S. Varadhan, *Diffusion processes in a small time interval,* Comm. Pure Appl. Math. 20 (1967), 659–685.

[37] S.R.S. Varadhan, *Large Deviations and Applications,* CBMS 46, SIAM, Philadelphia, 1984.

[38] J.B. Walsh, *An introduction to stochastic partial differential equations,* Ecole d'Ete de Probabilite de Saint-Flour XIV (1984), P.L. Hennequin editor, Lecture Notes in Mathematics 1180 , 265–439.

[39] F.-Y. Wang, *Harnack Inequality and Applications for Stochastic Generalized Porous Media equations,* Ann. Probab. 35(2007), 1333-1350.

[40] L. Wu, *On large deviations for moving average processes,* In Probability, Finance and Insurance, pp.15-49, the proceeding of a Workshop at the University of Hong-Kong (15-17 July 2002), Eds: T.L. Lai, H.L. Yang and S.P. Yung. World Scientific 2004, Singapour.

[41] E. Zeidler, *Nonlinear Functional Analysis and its Applications, II/B, Nonlinear Monotone Operators,* Springer-Verlag, New York: 1990.

[42] X. Zhang, *On Stochastic evolution equations with non-Lipschitz coefficients,* Preprint.